PATENT

Attorney Docket No. 67900/DJB/RMS/DAV

HUMAN KINESINS AND METHODS OF PRODUCING AND PURIFYING HUMAN KINESINS

FIELD OF THE INVENTION

This invention relates to the production and purification of human kinesins, preferably functional, using prokaryotic systems and to human kinesins isolated from bacterial systems.

BACKGROUND OF THE INVENTION

Cancer is the second-leading cause of death in industrialized nations. Effective therapeutics include the taxanes and vinca alkyloids, agents which act on microtubules. Microtubules are the primary structural element of the mitotic spindle. The mitotic spindle is responsible for distribution of replicate copies of the genome to each of the two daughter cells that result from cell division. It is presumed that it is the disruption of the mitotic spindle by these drugs that results in inhibition of cancer cell division, and also induction of cancer cell death. However, microtubules also form other types of cellular structures, including tracks for intracellular transport in nerve processes. Therefore, the taxanes have side effects that limit their usefulness.

Mitotic kinesins are enzymes essential for assembly and function of the mitotic spindle, but are not generally part of other microtubule structures, such as nerve processes. Mitotic kinesins play essential roles during all phases of mitosis. These enzymes are "molecular motors" that translate energy released by hydrolysis of ATP into mechanical force which drives the directional movement of cellular cargoes along microtubules. The catalytic domain sufficient for this task is a compact structure of approximately 340 amino acids. During mitosis, kinesins organize microtubules into the bipolar structure that is the mitotic spindle. Kinesins mediate movement of chromosomes along spindle microtubules, as well

as structural changes in the mitotic spindle associated with specific phases of mitosis. Experimental perturbation of mitotic kinesin function causes malformation or dysfunction of the mitotic spindle, frequently resulting in cell cycle arrest. From both the biological and enzymatic perspectives, these enzymes are attractive targets for the discovery and development of novel anti-mitotic chemotherapeutics.

A number of kinesins have been described in the art. However, there still exists a need for kinesins which can be easily produced in large quantities. In particular, human mitotic kinesins isolated and purified from a bacterial source are desirable.

Among the kinesins which have been identified is chromokinesin. Chromokinesin is a kinesin localized to mitotic chromatin and contributes to prometaphase chromosome alignment; it is expressed primarily in proliferating tissues and is enriched in mitotic compared to interphase cells. Perturbation of a Xenopus chromokinesin causes gross defects in mitotic spindle formation, including dissociation of chromosomes from spindle microtubules, multipolar spindles, misaligned chromosomes and failure of cytokinesis. Cloning of chicken (Wang and Adler, J. Cell Biol., 128:761-8 (1995)) and human (Oh, et al., direct GenBank submission without corresponding publication, submitted June 11, 1998 by Molecular Biology, Institute for Medical Sciences, San5 Wonchon Paldal, Suwon, Kyongki 442-749, Korea) chromokinesin homologs have been reported. The mouse homolog of chromokinesin, Kif4, has been expressed in Sf9 cells (bacculovirus vector) and has been reported to have motility and ATPase activity (Sekine, et al., J. Cell Biol., 127-187-201 (1994)). This same study speculated that Kif4 may participate in the transport of membraneous organelles in neuronal and other cell types.

Another kinesin reported to be associated with chromosomes is Kid. Kid is reported as unrelated to other known kinesins. The C-terminal 260 amino acids of Kid expressed in bacteria and purified reportedly binds directly to DNA in vitro. The same study reported that when fused to a transcriptional activation domain and co-transfected into mammalian cells with a reporter construct this domain can stimulate expression from the promoter on the co-transfected construct in living cells. Tokai, et al., EMBO J., 15(3):457-467 (1996). This study further reports that the amino-terminal 470 amino acids of Kid, which includes the motor domain, has been expressed in bacteria fused to glutathione-S-transferase, binds to microtubues and exhibits microtubule-stimulated ATPase activity. Kid is expressed in all

human cell lines that have been examined, and is most abundant in adult human speen, thymus and testis as well as fetal liver and kidney. In cultured human cells, Kid is reportedly found associated with chromatin throughout mitosis, showing some enrichment at kinetochores.

Another mitotic kinesin of interest is MKLP1 which localizes to microtubules of the spindle midzone throughout mitosis. In vitro MKLP1 can slide antiparallel microtubules relative to each other. Microinjection of antibody directed against MKLP1 into mammalian cells induces mitotic arrest with subtle defects in microtubule organization. Genetic data from both Drosophila and C. elegans clearly show that MKLP1 homologues are required for organization of the interzonal microtubules of the anaphase spindle and for formation of a functional contractile ring. MKLP1 associates with a kinase of the polo family in both Drosophila and mammals. Cloning of human (Nislow, et al., Nature, 359:543-7 (1992)), hamster (Kuriyama, et al., J. Cell Sci., 107(Pt 12):3485-99 (1994)), Drosophila (Adams, et al., Gene Dev., 12:1483-94 (1998)), and C. elegans (Raich, et al., Mol. Biol. Cell, 9:2037-49 (1998)) homologs of MKLP1 have been reported. Nislow, et al., supra, reported on expressed full-length human MKLP1 in bacteria, however there was relatively poor expression, and the polypeptide was not purified. Using this crude bacterial lysate, microtubule bundling and sliding activity were reported on. Kuriyama, et al., supra, reported on expressed hamster MKLP1 in Sf9 cells (baculovirus vector), but the protein was not purified.

KSP is also of interest. KSP belongs to an evolutionarily conserved kinesin subfamily of plus end-directed microtubule motors that assemble into bipolar homotetramers consisting of antiparallel homodimers. During mitosis KSP associates with microtubules of the mitotic spindle. Microinjection of antibody directed against KSP into human cells prevents spindle pole separation during prometaphase, giving rise to monopolar spindles and causing mitotic arrest. KSP and related kinesins bundle antiparallel microtubules and slide them relative to one another, thus forcing the two spindle poles apart. KSP may also mediate in anaphase B spindle elongation and focussing of microtubules at the spindle pole. Cloning of human (Blangy, et al., Cell, 83:1159-69 (1995); Whitehead, et al., direct GenBank submission without corresponding publication, submitted September 29, 1995 by Medical Biochemistry, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta TN 4N1, Canada), Drosophila (Heck, et al., J Cell Biol, 123:665-79 (1993)) and Xenopus (Le Guellec, et al., Mol. Cell Biol.,

11(6):3395-8 (1991)) homologs of KSP have been reported. Drosophila KLP61F/KRP130 has reportedly been purified in native form (Cole, et al., J. Biol. Chem., 269(37):22913-6 (1994)), expressed in E. coli, (Barton, et al., Mol. Biol. Cell, 6:1563-74 (1995)) and reported to have motility and ATPase activities (Cole, et al., supra; Barton, et al., supra). Xenopus Eg5 was expressed in E.coli and reported to possess motility activity (Sawin, et al., Nature, 359:540-3 (1992); Lockhart and Cross, Biochemistry, 35(7):2365-73 (1996); Crevel, et al, J. Mol. Biol., 273:160-170 (1997) and ATPase activity (Lockhart and Cross, supra; Crevel et al., supra).

CENP-E, also of interest, is a plus end-directed microtubule motor essential for achieving metaphase chromosome alignment. CENP-E accumulates during interphase and is degraded following completion of mitosis. Microinjection of antibody directed against CENP-E or overexpression of a dominant negative mutant of CENP-E causes mitotic arrest with prometaphase chromosomes scattered on a bipolar spindle. The tail domain of CENP-E mediates localization to kinetochores and also interacts with the mitotic checkpoint kinase hBubR1. CENP-E also associates with active forms of MAP kinase. Cloning of human (Yen, et al., Nature, 359(6395):536-9 (1992)) CENP-E has been reported. In Thrower, et al., EMBO J., 14:918-26 (1995) partially purified native human CENP-E was reported on. Moreover, the study reported that CENP-E was a minus end-directed microtubule motor. Wood, et al., Cell, 91:357-66 (1997)) discloses expressed Xenopus CENP-E in *E. coli* and that XCENP-E has motility as a plus end directed motor in vitro.

The kinesin MCAK has also been identified. During anaphase A disjoined sister chromatids migrate poleward. This poleward movement is driven by kinetochores and is accompanied by the depolymerization of microtubules attached to the migrating chromatids. The kinesin MCAK plays an important role in this motility and may promote disassembly of microtubules attached to kinetochores. MCAK localizes to kinetochores of mitotic chromosomes. MCAK belongs to small and unique subfamily of kinesins (Kin I) that destabilize microtubule ends. Overexpression of a dominant negative MCAK mutant or antisense inhibition of MCAK expression causes chromosomes to lag during anaphase. Genes for the Xenopus (Walczak, et al., Cell, 84:37-47 (1996), hamster (Wordeman and Mitchison, J. Cell Biol., 128:95-104 (1995) and human (Kim, et al., Biochim. Biophys. Acta., 1359:181-6 (1997)) homologs of MCAK have been cloned and characterized. Kim, et al., supra, also described

mRNA expression patterns of the endogenous gene in human cells and tissues, but did not describe exogenous expression.

Other mitotic kinesins of interest include HSET and Kif15. However, it is understood, as described above, all kinesins are of interest.

The kinesin superfamily further encompasses a number of families that exhibit non-mitotic motor functions, e.g., vesicle and organelle transport. These proteins are ATP dependent, and have plus end-directed microtubule motor activity involved in fast anterograde organelle transport in neurons. Anterograde transport is a directional transport, typically of mitochondria, other organelles and vesicles, from the cell body to the tip of the axon. Moreover, some non-mitotic kinesins are involved in "slow" transport.

Among the kinesins associated with neurons is the Kif2 family of kinesins. Cloning of mouse (Aizawa, et al., Genes Dev., 12:1483-94 (1992)), Xenopus (Walczak, et al., supra), and human (Debernardi, et al., Genomics, 42:67-73 (1997)) Kin2 homologs have been reported. Mouse Kif2 (Noda, et al., J. Cell iol., 129:157-67 (1995)) was reportedly expressed in Sf9 cells (bacculovirus vector) and was reported to have motility activity. Xenopus Kif2 (Desai, et al., Cell, 96:69-78 (1999)) was expressed in Sf9 cells (bacculovirus vector) and microtubule depolymerization activity was reported.

Cloning of human Kif1A (ATSV) has been reported (Furlong, et al., Genomics, 33(3):421-29 (1996)). The mouse homolog was expressed in bacculovirus and characterized biochemically (Okada, et al., Cell, 81:769-80 (1995)), and a mouse Kif1A/KHC hybrid (formed for stability) was also expressed in E. coli and reportedly had activity in a motility assay (Okada and Hirokawa, Science, 283:1152-7 (1999)).

The human form of KHC (Kinesin Heavy Chain) has been cloned (Navone, et al., J. Cell Biol, 117:1263-75 (1992)). Fujiwara, et al., Biophys J., 69:1563-8 (1995) reportedly expressed human KHC fragment 1-349 in *E.coli* and conducted structural studies on the purified protein. Vale, et al., Nature, 380:451-3 (1996) reportedly expressed human KHC fragment 1-560 in *E. coli* and purified it by phosphocellulose and Mono-Q chromatography. KHC additionally reportedly had motility activity.

Functional studies of enzymes, including but not limited to high-throughput screening for small molecule inhibitors, require significant amounts of active protein. Therefore, it is an object of this invention to provide systems to express kinesins in high quantities. It is further an object of this invention to provide methods for expression and purification of kinesins. It is further an object to provide kinesins which are unglycosylated.

SUMMARY OF THE INVENTION

Accordingly, the present invention provides kinesins which are produced from prokaryotes. In a preferred embodiment, bacterial systems are utilized. Bacterial expression provides the most economical means of obtaining substantial quantities of kinesins without a concern for copurifying a contaminating activity from the expression host since bacteria do not harbor kinesin like proteins.

In one aspect, the invention provides a method of producing a human mitotic kinesin protein comprising a motor domain. The method comprises expressing a nucleic acid comprising a nucleic acid encoding a human mitotic kinesin protein comprising a motor domain in a bacterial cell and substantially purifying said human mitotic kinesin protein.

In another aspect, a method is provided for producing a human kinesin protein comprising a motor domain and at least two epitope tags. The method comprises expressing a nucleic acid encoding a human kinesin protein comprising a motor domain and at least two epitope tags in a prokaryote and substantially purifying said human kinesin protein.

In a further aspect, a method is provided for producing a kinesin protein comprising a motor domain. The method comprises expressing a nucleic acid encoding a kinesin protein comprising a motor domain in a prokaryote and substantially purifying said kinesin protein, wherein said kinesin is selected from the group consisting of Kin2, chromokinesin, Kif1A and MKLP1. It is understood that unless a particular species is named, the term "kinesin" includes homologs thereof which may have different nomenclature among species. For example, the human homolog of Kif1A is termed ATSV, the human homologue of Xenopus Eg5 is termed KSP, and human HSET corresponds to Chinese hamster CHO2.

Also provided herein is a substantially pure unglycosylated human mitotic kinesin protein

comprising a motor domain. A substantially pure unglycosylated human kinesin protein comprising a motor domain and at least two epitope tags is also provided. Additionally, a substantially pure unglycosylated kinesin protein comprising a motor domain, wherein said kinesin is selected from the group consisting of Kin2, chromokinesin, Kif1A and MKLP1 is provided.

In one embodiment a prokaryote comprising a nucleic acid comprising a nucleic acid encoding a kinesin selected from the group consisting of chromokinesin, Kin2, and Kif1A is provided. In a further embodiment, a prokaryote comprising a nucleic acid comprising a nucleic acid encoding a human kinesin selected from the group consisting of chromokinesin, Kin2, Kif1A, KSP, CENP-E, MCAK, HSET and Kif15 is provided.

The proteins provided herein can be used in assays provided herein to determine binding properties and modulators of biological activity.

In a further embodiment, provided herein is a substantially purified unglycosylated peptide selected from the group consisting of K335, Q475, D679, FL1, P166, H195, FL2, E433, R494, E658, L360, K491, S553, M329, T340, S405, V465, T488, M1, M2, M3, M4, M5, M6, FL3, A2N370, A2M511, K519, E152.2, Q151.2, Q353 and M472.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the objects of this invention, methods of producing kinesins are provided herein. In a preferred embodiment, the kinesins are produced from a prokaryote. In a preferred embodiment, the prokaryote is a bacterial cell. Bacterial expression offers several advantages over systems previously utilized, such as, for example, Bacculovirus. The yield of protein is higher, the cost of the expression setup is lower, and creation of alternative expression vectors is easier. The concern of copurifying a contaminating activity from the expression host is also eliminated since bacteria, in contrast to the bacculovirus expression system, do not have kinesin like proteins.

Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli*

strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Preferred bacterial strains include *E. coli* BL21 (DE3), BL21 (DE3), pLysS, BL21 (DE3) pLysE.

The kinesins that are produced by the methods herein each comprise a molecular motor domain. Therefore, in one embodiment the kinesin is a full length protein. In another embodiment the kinesin is a kinesin protein comprising a molecular motor domain. A molecular motor protein is a cytoskeletal molecule that utilizes chemical energy to produce mechanical force, and drives the motile properties of the cytoskeleton. The molecular motor domain is usually about 35-45% identical among all kinesin superfamily members, and is approximately 340 amino acids. Vale and Kreis, 1993, GUIDEBOOK TO THE CYTOSKELETAL AND MOTOR PROTEINS New York: Oxford University Press; Goldstein, 1993, Ann. Rev. Genetics 27: 319-351; Mooseker and Cheney, 1995, Annu. Rev. Cell Biol. 11: 633-675; Burridge et al., 1996, Ann. Rev. Cell Dev. Biol. 12: 463-519.

In one embodiment, the kinesin can be from any species. Thus, unless otherwise specified, kinesin includes homologs thereof. The kinesins therefore include those from Xenopus, Drosophila and other insects, plants, fungi and mammalian cells, with rodents (mice, rats, hamsters, guinea pigs and gerbils being preferred), primates and humans being preferred. In a preferred embodiment, the kinesin is selected from the group consisting of chromokinesin, Kin2, Kif1A, and MKLP1. Preferably Kif1A is expressed as an individual kinesin, i.e., it excludes fusion forms to other kinesins.

In another embodiment, the kinesin is a human kinesin. In a preferred embodiment, the human kinesin is selected from the group consisting of chromokinesin, Kin2, Kif1A, MKLP1, KSP, CENP-E, MCAK, KHC, HSET and Kif15.

In one embodiment, the kinesin protein is a mitotic kinesin protein. In one embodiment, the mitotic kinesin protein is selected from the group consisting of chromokinesin, MKLP1, KSP, CENP-E and MCAK. In a preferred embodiment, the mitotic kinesin protein is a human mitotic kinesin protein.

In another embodiment, the kinesin protein is a non-mitotic kinesin protein. In a preferred embodiment, the non-mitotic kinesin protein is selected from the group consisting of KHC, Kin2 and Kif1A. In a preferred embodiment, the non-mitotic kinesin protein is a human non-mitotic kinesin protein.

In a particularly preferred embodiment, the human kinesin protein is selected from the group consisting of chromokinesin, KSP, CENP-E, MCAK, Kin2 and Kif1A. In another particularly preferred embodiment, the kinesin protein is selected from the group consisting of chromokinesin, Kif1A, MKLP1 and Kin2, with chromokinesin and Kin2 being most preferred. It is understood that the groups provided herein necessarily describe groups or individuals within them. For example, the group consisting of KSP, CENP-E, MCAK, Kin2 and Kif1A describes a group consisting of KSP, CENP-E, MCAK and Kin2, or CENP-E as an individual kinesin protein, etc.

In another embodiment, the kinesin protein is a peptide selected from the group consisting of K335, Q475, D679, FL1, P166, H195, FL2, E433, R494, E658, L360, K491, S553, M329, T340, S405, V465, T488, M1, M2, M3, M4, M5, M6, FL3, A2N370, A2M511, K519, E152.2, Q151.2, Q353 and M472. Similarly, it is understood that this group explicitly includes the group of M1, M2, and M6 or K335 and K491, etc.

In one embodiment, the kinesin proteins provided herein have glycosylation patterns which differs from their native form. In a preferred embodiment, the kinesin proteins provided herein are unglycosylated. In a preferred embodiment, the kinesin proteins are expressed in prokaryotes, preferably bacteria, and most preferably E.coli, and are thus unglycosylated. However, it is understood that at least in one embodiment an altered native glycosylation pattern can be obtained by a variety of techniques. Removal of carbohydrate moieties present on the kinesin protein may further be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

In another aspect, the kinesins provided herein may have phosphorylation or farnesylation patterns which differ from their native form. In one embodiment, a kinesin is provided which substantially lacks phosphorylation, farneslation and glycosylation.

In one embodiment provided herein, the kinesin protein has at least one and preferably at least two epitope tags. An example of such a chimeric molecule comprises a kinesin protein fused to an epitope tag sequence or a Fc region of an immunoglobulin. The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a kinesin protein fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least five amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 6 and 20 amino acid residues). In one embodiment, such a chimeric molecule comprises a fusion of the kinesin protein with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the kinesin protein. The presence of such epitope-tagged forms of the kinesin protein can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the kinesin protein to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In a preferred embodiment, the kinesin protein comprises an N-terminal T7 epitope tag and a C-terminus myc-epitope and 6-His tag.

In an alternative embodiment, the chimeric molecule may comprise a fusion of the kinesin protein with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a Kinesin protein in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG-1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

Additionally, as recognized by the skilled artisan and as will be further apparent below, labels of various sorts may be utilized in the invention. A "label" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P, fluorescent dyes, electron-dense reagents, enzymes, biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available. Labels are also described further below.

In a preferred embodiment, a method provided herein includes purifying said kinesin protein. The terms "isolated" "purified" or "biologically pure" refer to material that is substantially or essentially free from components which normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure. In a preferred embodiment, a protein is considered pure wherein it is determined that there is no contaminating activity.

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding the kinesin protein may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression.

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides which have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences and as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res. 19*:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem. 260*:2605-2608 (1985); Cassol *et al.*, 1992; Rossolini *et al.*, *Mol. Cell. Probes 8*:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The kinesin protein may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the kinesin-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a

variety of bacteria. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

Expression and cloning vectors usually contain a promoter operably linked to the kinesin-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding kinesin protein.

The host cells are transformed with the nucleic acids as described herein for kinesin protein production and cultured in nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The preferred embodiments are demonstrated in the examples below.

Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Bacteria are grown according to standard procedures in the art. Preferably fresh bacteria cells are used for isolation of protein.

The preferred embodiments for each of the steps of production and purification are further described below in the examples. In particular, preferred lysis, wash and elution buffers are provided. In a preferred embodiment, purification over a Ni-NTA resin leads to a high degree of purification in a single step.

Preferably, the kinesins provided herein as compositions or produced from the methods provided herein have at least one activity of a kinesin protein as further defined below. Preferably the activity is the ability to hydrolyze ATP in a manner stimulated by microtubules.

While it is preferable to produce the kinesins herein in prokaryotic systems, in one aspect, the kinesins herein are produced in eukaryotic systems. In each case, the kinesin is expressed recombinantly. Previous work provided a limited number of kinesin homologs recombinantly, however, herein, each homolog, preferably the human homolog, is expressed recombinantly. For example, methods for expressing human Kin2 in a recombinant system are provided herein. In a preferred embodiment, a vector comprising a human Kin2 sequence is expressed in a eukaryotic cell, and the Kin2 is purified. Similarly, in one embodiment, human chromokinesin, HSET, Kif15, MCAK, Kif1A, MKLP1, CENP-E, KHC or KSP is expressed in a eukaryotic cell. In a preferred embodiment, the eukaryotic cell works in conjunction with a baculovirus system, such as Sf9 cell. The kinesins provided produced by such systems are also provided herein.

In one aspect the specific coding sequences as published and known in the art which encode the kinesin proteins are utilized. However, in an alternative embodiment, a substantially identical sequence encoding a kinesin protein is utilized. The term "substantially identical" in the context of two nucleic acids or polypeptides refers to the residues in the two sequences that have at least 80% identity when aligned for maximum correspondence as measured using one of the following algorithms. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2: 482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP

uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5: 151-153 (1989). The program can align up to 300 sequences of a maximum length of 5,000. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster can then be aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences can be aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program can also be used to plot a dendogram or tree representation of clustering relationships. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison, e.g., the motor domain. In one example, kinesin proteins were compared to other kinesin protein superfamily sequences using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps.

Another example of algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul *et al., J. Mol. Biol.* 215: 403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90: 5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a nucleic acid if the smallest sum probability in a comparison of the test nucleic acid to the nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001. Where the test nucleic acid encodes a kinesin protein, it is considered similar to a specified kinesin nucleic acid if the comparison results in a smallest sum probability of less than about 0.5, and more preferably less than about 0.2.

An indication that two polypeptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a polypeptide is substantially identical to a second polypeptide, for example, where the two peptides differ only by a conservative substitution. An indication that two nucleic acid sequences are substantially identical is that the polypeptide which the first nucleic acid encodes is immunologically cross reactive with the polypeptide encoded by the second nucleic acid. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions.

The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA. The term "stringent conditions" refers to conditions under which a probe will hybridize to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium (as the target sequences are generally present in excess, at T_m, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion,

typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

Once expressed and purified if necessary, the kinesin proteins and nucleic acids are useful in a number of applications.

In a preferred embodiment, the kinesin proteins or cells containing the native or modified kinesin proteins are used in screening assays. Production of these important motor proteins in large quantities permits the design of drug screening assays for compounds that modulate kinesin activity.

Screens may be designed to first find candidate agents that can bind to kinesin proteins, and then these agents may be used in assays that evaluate the ability of the candidate agent to modulate kinesin activity. Thus, as will be appreciated by those in the art, there are a number of different assays which may be run; binding assays and activity assays.

Thus, in a preferred embodiment, the methods comprise combining a kinesin protein and a candidate bioactive agent, and determining the binding of the candidate agent to the kinesin protein. Preferred embodiments utilize a human kinesin protein, although other homologs may be used. In a preferred embodiment, the kinesin is unglycosylated or has at least two epitope tags as described herein.

The term "candidate bioactive agent" as used herein describes any molecule, e.g., protein, oligopeptide, small organic molecule, polysaccharide, polynucleotide, etc., with the capability of directly or indirectly altering the bioactivity of kinesin. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary

for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification to produce structural analogs.

In a preferred embodiment, the candidate bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides. The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homo-phenylalanine, citrulline and noreleucine are considered amino acids for the purposes of the invention. "Amino acid" also includes imino acid residues such as proline and hydroxyproline. The side chains may be in either the (R) or the (S) configuration. In the preferred embodiment, the amino acids are in the (S) or L-configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard in vivo degradations.

In a preferred embodiment, the candidate bioactive agents are naturally occurring proteins or fragments of naturally occurring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of procaryotic and eucaryotic proteins may be made for screening against

kinesin. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the latter being preferred, and human proteins being especially preferred.

In a preferred embodiment, the candidate bioactive agents are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized candidate bioactive proteinaceous agents.

In one embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or are selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

In a preferred embodiment, the candidate bioactive agents are nucleic acids. By "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramide (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579 (1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., Chemica Scripta 26:141 91986)),

phosphorothioate (Mag et al., Nucleic Acids Res. 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu et al., J. Am. Chem. Soc. 111:2321 (1989), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc. 114:1895 (1992); Meier et al., Chem. Int. Ed. Engl. 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson et al., Nature 380:207 (1996), all of which are incorporated by reference). Other analog nucleic acids include those with positive backbones (Denpcy et al., Proc. Natl. Acad. Sci. USA 92:6097 (1995); non-ionic backbones (U.S. Patent Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowshi et al., Angew. Chem. Intl. Ed. English 30:423 (1991); Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); Letsinger et al., Nucleoside & Nucleotide 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., Bioorganic & Medicinal Chem. Lett. 4:395 (1994); Jeffs et al., J. Biomolecular NMR 34:17 (1994); Tetrahedron Lett. 37:743 (1996)) and non-ribose backbones, including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins et al., Chem. Soc. Rev. (1995) pp169-176). Several nucleic acid analogs are described in Rawls, C & E News June 2, 1997 page 35. All of these references are hereby expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be done to facilitate the addition of additional moieties such as labels, or to increase the stability and half-life of such molecules in physiological environments. In addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made. The nucleic acids may be single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xathanine hypoxathanine, isocytosine, isoguanine, etc.

As described above generally for proteins, nucleic acid candidate bioactive agents may be naturally occurring nucleic acids, random nucleic acids, or "biased" random nucleic acids.

For example, digests of procaryotic or eucaryotic genomes may be used as is outlined above for proteins.

In a preferred embodiment, the candidate bioactive agents are organic chemical moieties, a wide variety of which are available in the literature.

The assays described utilize kinesin proteins as defined herein. In one embodiment, portions of kinesin proteins are utilized, in a preferred embodiment, portions having kinesin activity are used. In addition, the assays described herein may utilize either isolated kinesin proteins or cells comprising the kinesin proteins.

In one of the embodiments of the methods provided herein, the kinesin protein or the candidate agent is non-diffusably bound to an insoluble support having isolated sample receiving areas (e.g. a microtiter plate, an array, etc.). The insoluble supports may be made of any composition to which the compositions can be bound, is readily separated from soluble material, and is otherwise compatible with the overall method of screening. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports include microtiter plates, arrays, membranes and beads. These are typically made of glass, plastic (e.g., polystyrene), polysaccharides, nylon or nitrocellulose, teflon™, etc. Microtiter plates and arrays are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples. In some cases magnetic beads and the like are included. The particular manner of binding of the composition is not crucial so long as it is compatible with the reagents and overall methods of the invention, maintains the activity of the composition and is nondiffusable. Preferred methods of binding include the use of antibodies (which do not sterically block either the ligand binding site or activation sequence when the protein is bound to the support), direct binding to "sticky" or ionic supports, chemical crosslinking, the synthesis of the protein or agent on the surface, etc. Following binding of the protein or agent, excess unbound material is removed by washing. The sample receiving areas may then be blocked through incubation with bovine serum albumin (BSA), casein or other innocuous protein or other moiety. Also included in this invention are screening assays wherein solid supports are not used. Solution based assays are further described below.

In a preferred embodiment, the kinesin protein is bound to the support, and a candidate bioactive agent is added to the assay. Alternatively, the candidate agent is bound to the support and the kinesin protein is added. Novel binding agents include specific antibodies, non-natural binding agents identified in screens of chemical libraries, peptide analogs, etc. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, functional assays (phosphorylation assays, etc.) and the like.

The determination of the binding of the candidate bioactive agent to a kinesin protein may be done in a number of ways. In a preferred embodiment, the candidate bioactive agent is labelled, and binding determined directly. For example, this may be done by attaching all or a portion of a kinesin protein to a solid support, adding a labelled candidate agent (for example a fluorescent label), washing off excess reagent, and determining whether the label is present on the solid support. Various blocking and washing steps may be utilized as is known in the art.

By "labeled" herein is meant that the compound is either directly or indirectly labeled with a label which provides a detectable signal, e.g. radioisotope, fluorescers, enzyme, antibodies, particles such as magnetic particles, chemiluminescers, or specific binding molecules, etc. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule which provides for detection, in accordance with known procedures, as outlined above. The label can directly or indirectly provide a detectable signal.

In some embodiments, only one of the components is labeled. For example, the proteins (or proteinaceous candidate agents) may be labeled at tyrosine positions using ^{125}I, or with fluorophores. Alternatively, more than one component may be labeled with different labels; using ^{125}I for the proteins, for example, and a fluorophor for the candidate agents.

In a preferred embodiment, the binding of the candidate bioactive agent is determined through the use of competitive binding assays. In this embodiment, the competitor is a binding moiety known to bind to the target molecule (i.e. kinesin), such as an antibody,

peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding as between the bioactive agent and the binding moiety, with the binding moiety displacing the bioactive agent. This assay can be used to determine candidate agents which interfere with binding between kinesin proteins and, for example, a microtubule.

In one embodiment, the candidate bioactive agent is labeled. Either the candidate bioactive agent, or the competitor, or both, is added first to the protein for a time sufficient to allow binding, if present. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high through put screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

In a preferred embodiment, the competitor is added first, followed by the candidate bioactive agent. Displacement of the competitor is an indication that the candidate bioactive agent is binding to the kinesin protein and thus is capable of binding to, and potentially modulating, the activity of the kinesin protein. In this embodiment, either component can be labeled. Thus, for example, if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the candidate bioactive agent is labeled, the presence of the label on the support indicates displacement.

In an alternative embodiment, the candidate bioactive agent is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate that the bioactive agent is bound to the kinesin protein with a higher affinity. Thus, if the candidate bioactive agent is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate that the candidate agent is capable of binding to the kinesin protein.

In a preferred embodiment, the methods comprise differential screening to identity bioactive agents that are capable of modulating the activity of the kinesin proteins. In this embodiment, the methods comprise combining a kinesin protein and a competitor in a first sample. A second sample comprises a candidate bioactive agent, a kinesin protein and a

competitor. The binding of the competitor is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the kinesin protein and potentially modulating its activity. That is, if the binding of the competitor is different in the second sample relative to the first sample, the agent is capable of binding to the kinesin protein.

Alternatively, a preferred embodiment utilizes differential screening to identify drug candidates that bind to the native kinesin protein, but cannot bind to modified kinesin proteins. The structure of the kinesin protein may be modeled, and used in rational drug design to synthesize agents that interact with that site. Drug candidates that affect kinesin bioactivity are also identified by screening drugs for the ability to either enhance or reduce the activity of the protein.

Positive controls and negative controls may be used in the assays. Preferably all control and test samples are performed in at least triplicate to obtain statistically significant results. Incubation of all samples is for a time sufficient for the binding of the agent to the protein. Following incubation, all samples are washed free of non-specifically bound material and the amount of bound, generally labeled agent determined. For example, where a radiolabel is employed, the samples may be counted in a scintillation counter to determine the amount of bound compound.

A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding.

Screening for agents that modulate the activity of kinesin protein may also be done. In a preferred embodiment, methods for screening for a bioactive agent capable of modulating the activity of kinesin protein comprise the steps of adding a candidate bioactive agent to a sample of kinesin protein, as above, and determining an alteration in the biological activity of kinesin protein. "Modulating the activity of kinesin protein" includes an increase in activity, a decrease in activity, or a change in the type or kind of activity present. Thus, in this

embodiment, the candidate agent should both bind to kinesin protein (although this may not be necessary), and alter its biological or biochemical activity as defined herein. The methods include both in vitro screening methods, as are generally outlined above, and in vivo screening of cells for alterations in the presence, distribution, activity or amount of kinesin protein.

Thus, in this embodiment, the methods comprise combining a kinesin protein sample and a candidate bioactive agent, and evaluating the effect on motor activity. By "kinesin protein activity" or grammatical equivalents herein is meant one of kinesin protein's biological activities, including, but not limited to, its ability to affect ATP hydrolyzation. Other activities include microtubule binding, gliding, polymerazation/depolymerazation (effects on microtubule dynamics), binding to other proteins of the spindle, binding to proteins involved in cell-cycle control, or serving as a substrate to other enzymes, such as kinases or proteases and specific kinesin cellular activities such as chromosome congregation, axonal transport, etc.

Methods of performing motility assays are well known to those of skill in the art (*see, e.g.*, Hall, *et al.* (1996), *Biophys. J.*, 71: 3467-3476, Turner *et al.*, 1996, *Anal. Biochem.* 242 (1):20-5; Gittes *et al.*, 1996, *Biophys. J.* 70(1): 418-29; Shirakawa *et al.*, 1995, *J. Exp. Biol.* 198: 1809-15; Winkelmann *et al.*, 1995, *Biophys. J.* 68: 2444-53; Winkelmann *et al.*, 1995, *Biophys. J.* 68: 72S, and the like).

In addition to the assays described above, methods known in the art for determining ATPase activity can be used. Preferably, solution based assays are utilized. Alternatively, conventional methods are used. For example, P_i release from kinesin can be quantified. In one preferred embodiment, the ATPase activity assay utilizes 0.3 M PCA (perchloric acid) and malachite green reagent (8.27 mM sodium molybdate II, 0.33 mM malachite green oxalate, and 0.8 mM Triton X-100). To perform the assay, 10 µL of reaction is quenched in 90 µL of cold 0.3 M PCA. Phosphate standards are used so data can be converted to mM inorganic phosphate released. When all reactions and standards have been quenched in PCA, 100 µL of malachite green reagent is added to the to relevant wells in *e.g.*, a microtiter plate. The mixture is developed for 10-15 minutes and the plate is read at an absorbance of 650 nm. If phosphate standards were used, absorbance readings can be converted to mM

P_i and plotted over time. Additionally, ATPase assays known in the art include the luciferase assay.

In a preferred embodiment, the activity of the kinesin protein is decreased or increased, with a decrease being preferred. Modulation also includes changes such as the binding characteristics etc. Thus, bioactive agents that are antagonists are preferred in some embodiments, and bioactive agents that are agonists may be preferred in other embodiments.

The components provided herein for the assays provided herein may also be combined to form kits. The kits can be based on the use of the protein and/or the nucleic acid encoding the kinesin proteins.

In this way, bioactive agents are identified. Compounds with pharmacological activity are able to enhance or interfere with the activity of the kinesin protein. The compounds having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a host, as previously described. The agents may be administered in a variety of ways, orally, parenterally e.g., subcutaneously, intraperitoneally, intravascularly, etc. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100 wt.%.

The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils and fats. Stabilizing agents, wetting and emulsifying agents, salts for varying the osmotic pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents.

It is also understood that bioactive compounds may be used in the agricultural arena. For example, inhibitors of kinesins may eliminate fungi which adversely effect agricultural crops. Alternatively, inhibitors of kinesins may be useful in eliminating unwanted plants, i.e., weeds.

Thus, in one embodiment, methods of modulating motor activity in cells or organisms are provided. In one embodiment, the methods comprise administering to a cell an anti-kinesin antibody or other agent identified herein or by the methods provided herein, that reduces or eliminates the biological activity of the endogenous kinesin protein. Alternatively, the methods comprise administering to a cell or organism a recombinant nucleic acid encoding a kinesin protein or modulator including anti-sense nucleic acids.

In one embodiment, the kinesin proteins of the present invention may be used to generate polyclonal and monoclonal antibodies to kinesin proteins, which are useful as described herein. Similarly, the kinesin proteins can be coupled, using standard technology, to affinity chromatography columns. These columns may then be used to purify kinesin antibodies. In a preferred embodiment, the antibodies are generated to epitopes unique to the kinesin protein; that is, the antibodies show little or no cross-reactivity to other proteins. These antibodies find use in a number of applications. For example, the kinesin antibodies may be coupled to standard affinity chromatography columns and used to purify kinesin proteins as further described below. The antibodies may also be used as blocking polypeptides, as outlined above, since they will specifically bind to the kinesin protein.

The anti-kinesin protein antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the kinesin protein polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

The anti-kinesin protein antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described

by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the kinesin protein polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against kinesin protein. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or

enzyme-linked immunosorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is

truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

The anti-kinesin protein antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et

al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the kinesin protein, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random

assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

The anti-kinesin protein antibodies of the invention have various utilities. For example, anti-kinesin protein antibodies may be used in diagnostic assays for a kinesin protein, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic

assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-kinesin protein antibodies also are useful for the affinity purification of kinesin protein from recombinant cell culture or natural sources. In this process, the antibodies against kinesin protein are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the kinesin protein to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the kinesin protein, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the kinesin protein from the antibody.

The anti-kinesin protein antibodies may also be used in treatment. In one embodiment, the genes encoding the antibodies are provided, such that the antibodies bind to and modulate the kinesin protein within the cell.

All publications, sequences (those of known kinesins, those disclosed or referenced in publications cited herein, or those referenced herein by accession number) and patent applications cited in this specification are herein incorporated by reference as if each individual publication, sequence or patent application were specifically and individually indicated to be incorporated by reference in their entirety. Additionally, wherein accession numbers are provided for sequences herein, the related text in that database entry is also incorporated herein in its entirety.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and

modifications may be made thereto without departing from the spirit or scope of the appended claims.

EXAMPLES

EXAMPLE 1:

Bacterial Expression Constructs:

cDNA Cloning.

For all of the kinesin-related proteins provided herein as examples, the full-length human cDNA has been previously described (see Table I). We have cloned cDNAs for all examples by PCR using the primers and cDNA sources indicated on Table I, except for CENP-E which was obtained from Don Cleveland at the Ludwig Institute for Cancer Research, UCSD; see, Yen et al., Nature, 359(6395):536-9 (1992). The nucleotide numbering on Table I corresponds to the Genbank submission numbering scheme. The clones were all sequenced to confirm they were the same as the published genes, although some polymorphisms were present.

Table I: Cloning of Human kinesins

Huma kinesin	Published Seq: Accession #s & Publication Ref:	Primers for cDNA cloning: 5' primer 3' primer	Nucleo-tides Included	cDNA Source
Chromo-kinesin	AF071592 1165722*(GSDB – partial)*	RACE AP1 primer (Clontech) CCAAACAGGAAACAGTATCCAAGGCAACC	<1-193	Marathon-Ready HeLa (Clontech)
		TGCCCATCTCGTGAGAAAGC GCTTGACGGAGAGCATGCTG	76-1178	HeLa (Our prep)
		ATTGATTACCCAGTTATCGG TGATGACTCCAACTTCAGTG	1032-3326	HeLa (Our prep)
Kin-2	Y08319	GCCGAATACATCAAGCAATGGTAAC TCTGGGTATCCTTTAGCAGCAAATG	2-2088	Breast tumor (Invitrogen)
MKLP1	X67155 Nislow, et al. 1992	AGCCATGTTGTCAGCGAGAGCTAAG AGGGTCTCTCTGGCTTCTCAGTTTTAGG	73-2078	human placenta (Invitrogen)

KSP	U37426	CCTTGATTTTTTGGCGGGGACCGTC AAAGGTTGATCTGGGCTCGCAGAGG	66-3259	breast tumor (Invitrogen)
CENP-E	Z15005 Yen, et al. 1992			
MCAK	U63743 Kim, et al. 1997	GCGTTTCTCTTCCTTGCTGACTCTC AGAGGCTGGGTGTCAAACCAAACAG	22-2274	breast tumor (Invitrogen)
Kid	AB017430	GTCGCTGTCGGCTAAGCAAG CTTTGCCCCTGTGACTGTGC	101-1596	breast tumor (Invitrogen)
		CTGGATCCCAGCCGCGGGCGGCTCGACG GAG CTCTAGAGAGCAGCTGTCCATGCCCC	28-248	HeLa (our prep)
HSET	D14678 *(partial)*	GGGCTTGGTGCAAGAGCTTC CACCCCTCACCCGATACATAGAC	213-1624	HeLa (our prep)
ATSV	X90840	GGGCTCCCACTACTGCGAGG CTCCTCCTCGTTCACCTCCG	21-2311	WERI (our prep)

The sequences from the GenBank accession numbers from Table I and anywhere provided herein, are expressly and explicitly incorporated herein. Other preferred sequences include the following: HsATSV, GenBank accession number X90840; HsHSET/CHO2 partial, GenBank accession number D14678; HsKHC, GenBank accession number X65873; HsKid, GenBank accession number AB017430; and AnBimC, GenBank accession number M32075.

Expression Plasmid Vector Backbones:

Sub A2

pET23d (Novagen 69748-3) encodes a T7 epitope tag 5' of the polylinker cloning site and a 6-His tag 3' of the polylinker cloning site. We constructed pET23dmyc by inserting the annealed oligonucleotides described below into the XhoI site of pET23d. This creates coding sequence for the myc epitope tag in-frame with the 6-His tag.

Annealed oligonucleotides for pET23dmyc:

sense: TCGAGGGTACCGAGCAGAAGCTGATCAGCGAGGAGGACCTGA

antisense: TCGATCAGGTCCTCCTCGCTGATCAGCTTCTGCTCGGTACCC

pET15b (Novagen 69661-3) encodes a HIS tag 5' of the polylinker cloning site.

Subcloning of Genes into Expression Vectors:

Using the human kinesin clones obtained by the methods described above as a template, PCR was used to amplify portions of the coding sequence, and the PCR product was inserted into

the bacterial expression plasmids described above by restriction enzyme digest and ligation. Several constructs of different lengths were developed for each kinesin (see Table II, the column "Residues Included" describes the starting and ending amino acid in one-letter code and amino acid number). All of the resulting constructs encode the motor domain, and vary in the amount of flanking sequence. The PCR primers detailed on Table II are designed such that resulting constructs encode a protein with a C-terminal 6-His tag (for those constructs built into pET23d) or the combination myc-epitope/6-His tag (for those constructs built into pET23dmyc), or an N-temrinal 6-His tag (for those constructs built into pET15b). All constructs made using the pET23d or the pET23dmyc vector, except those cloned into the NcoI site, also encode a protein with an N-terminal T7 epitope tag.

Table II: Subcloning of Human kinesins into Bacterial Expression Plasmids:

Kinesin	Construct Name	5' primer / 3' primer	Residues Included	Cloning sites	Host Vector
Chromo-kinesin	K335	TAGCCATGGAAGAGGTGAAGGGAATTC	E2-	5': NcoI	pET23dmyc
		CCGCTCGAGTTTTCTTGCTCTGTC	K335	3': XhoI	
Chromo-kinesin	Q475	TAGAAGCTTGGAAGAGGTGAAGGG	E2-	5' Hind III	pET23dmyc
		TAGAAGCTTCTGGGTAATCAATTG	Q475	3': HindIII	
Chromo-kinesin	D679	TAGAAGCTTGGAAGAGGTGAAGGG	E2-	5' Hind III	pET23dmyc
		TAGAAGCTTGTCTCGTTCTTTTAAC	D679	3': HindIII	
Chromo-kinesin	FL1	TAGAAGCTTGGAAGAGGTGAAGGG	E2-	5' Hind III	pET23dmyc
		TAGAAGCTTGTGGGCCTCTTCTTCG	H1229	3': HindIII	
Kin2	P166	TACGGATCCCAAATTATGAAATTATG	P166-	5': BamHI	pET23dmyc
		TACAAGCTTAGCAGTTGGATCTACAGTC	A532	3': HindIII	
Kin2	H195	TACGGATCCATAGGATATGTGTGTGTG	H195-	5': BamHI	pET23dmyc
		TACAAGCTTAGCAGTTGGATCTACAGTC	A532	3': HindIII	
Kin2	FL2	CTCCATGGTAACATCTTTAAATGAAGATAATG	M1-	5': NcoI	pET23dmyc
		CTAAGCTTAAGGGCACGGGGTCTCTTCGGGTTG	L679	3': HindIII	
MKLP1	E433	ATCCATGGCGAGAGCTAAGACACCCCGGAAACC	A4-	5': NcoI	pET23dmyc
		ATGCGGCCGCTTCTTGAGTCACTTCCGCAAATCTC	E433	3': NotI	
MKLP1	R494	ATCCATGGCGAGAGCTAAGACACCCCGGAAACC	A4-	5': NcoI	pET23dmyc
		ATGCGGCCGCCCTTGGAAGTGTCTGCTCATCGTTG	R494	3': NotI	
MKLP1	E658	ATCCATGGCGAGAGCTAAGACACCCCGGAAACC	A4-	5': NcoI	pET23dmyc
		ATGCGGCCGCTTCAGTAACAATAGCCTTCAGTTG	E658	3': NotI	
KSP	L360	ATCCATGGCGTGCCAGCCAAATTCGTCTGCG	M1-	5': NcoI	pET23dmyc
		ATCTCGAGCAATATGTTCTTTGCTCTATGAGC	L360	3': XhoI	
KSP	K491	ATCCATGGCGTGCCAGCCAAATTCGTCTGCG	M1-	5': NcoI	pET23dmyc
		ATCTCGAGTTTCTCCTCAGTACTTTCCAAAGC	K491	3': XhoI	
KSP	S553	ATCCATGGCGTGCCAGCCAAATTCGTCTGCG	M1-	5': NcoI	pET23dmyc
		ATCTCGAGGCTGCCATCCTTAATTAATTCTTCC	S553	3': XhoI	
CENP-E	M329	CTGGATCCCGGCGGAGGAAGGAGCCGTGGCC	A2-	5': BamHI	pET23d
		CACTCGAGCATATATTTAGCAGTACTGGC	M329	3': XhoI	
CENP-E	T340	CTGGATCCCGGCGGAGGAAGGAGCCGTGGCC	A2-	5': BamHI	pET23d
		CACTCGAGAGTTGATACCTCATTAACATAAGGAG	T340	3': XhoI	
CENP-E	S405	CTGGATCCCGGCGGAGGAAGGAGCCGTGGCC	A2-	5': BamHI	pET23d
		CACTCGAGAGAAGAGGTCACCAGCATCCG	S405	3': XhoI	
CENP-E	V465	CTGGATCCCGGCGGAGGAAGGAGCCGTGGCC	A2-	5': BamHI	pET23d
		CACTCGAGGACAGATTCATCAATTTCTCG	V465	3': XhoI	
CENP-E	T488	CTGGATCCCGGCGGAGGAAGGAGCCGTGGCC	A2-	5': BamHI	pET23d
		CACTCGAGTGTTGCTGGATTCCATTCTATC	T488	3': XhoI	
MCAK	M1	CTGGATCCGGAGGAAATCATGTCTTGTGAAG	R189-	5': BamHI	pET23dmyc
		CACTCGAGTGGAATCAGCGCCCCGTTAGAG	P617	3': XhoI	
MCAK	M2	CTGGATCCCAAACTGGGAATTTGCCCGAATG	P228-	5': BamHI	pET23dmyc
		CACTCGAGTGGAATCAGCGCCCCGTTAGAG	P617	3': XhoI	

MCAK	M3	CTGGATCCACAGAATATGTGTCTGTGTTAGG	H257-	5': BamHI	pET23dmyc
		CACTCGAGTGGAATCAGCGCCCCGTTAGAG	P617	3': XhoI	
MCAK	M4	CTGGATCCGGAGGAAATCATGTCTTGTGAAG	R189-	5': BamHI	pET23dmyc
		CACTCGAGTGGTCCTTGCTGTATGATCTC	P660	3': XhoI	
MCAK	M5	CTGGATCCCAAACTGGGAATTTGCCCGAATG	P228-	5': BamHI	pET23dmyc
		CACTCGAGTGGTCCTTGCTGTATGATCTC	P660	3': XhoI	
MCAK	M6	CTGGATCCACAGAATATGTGTCTGTGTTAGG	H257-	5': BamHI	pET23dmyc
		CACTCGAGTGGTCCTTGCTGTATGATCTC	P660	3': XhoI	
MCAK	FL3	CTCCATGGACTCGTCGCTTCAGGCCCGC	M3-	5': NcoI	pET23dmyc
		CTCTCGAGCTGGGGCCGTTTCTTGCTGCTTATTT G	Q725	3': XhoI	
Kid	A2N370	CTGGATCCCAGCCGCGGGCGGCTCGACGCAG	A2-	5': BamHI	pET23dmyc
		CACTCGAGATTGATCACCTGCTTGGACCTG	N370	3': XhoI	
Kid	A2M511	CTGGATCCCAGCCGCGGGCGGCTCGACGCAG	A2-	5': BamHI	pET23dmyc
		CACTCGAGCATTGTGGGACAATGGTTCTC	M511	3': XhoI	
HSET	K519	TCGGATCCTTGGTGCAAGAGCTTCAG	L72-K519	5': BamHI	pET23dmyc
		CACTCGAGCTTCCTGTTGGCCTGAGC		3' XhoI	
HSET	E152.2	CATGCCATGGAACTCAAGGGCAAC	E152-K519	5': NcoI	pET23d
		CACTCGAGCTTCCTGTTGGCCTGAGC		3': XhoI	
HSET	Q151.3	GGATATCCATATGCAGGAACTCAAGGGCAAC	Q151-K519	5': NdeI	pET15b
		GCAGGATCCTCACTTCCTGTTGGCCTGAG		3': BamHI	
ATSV	Q353	CTGGATCCCCCGGGGCTTCGGTGAAGGTGGCG	G3-Q353	5': BamHI	pET23dmyc
		CACTCGAGCTGCTTGGCCCGGTCAGCATAC		3': XhoI	
ATSV	M472	CTGGATCCCCCGGGGCTTCGGTGAAGGTGGCG	G3-M472	5': BamHI	pET23dmyc
		CACTCGAGCATCTCGGCCAGCAGGGCTTC		3': XhoI	

The construct name, such as "Q475", is used herein to identify the construct initially identified by the "residues included" and the GenBank accession number provided herein. As noted in the procedures provided herein, the vector also supplies an initiation methionine and epitope tags. It is understood that when the construct is named in the context of a peptide, such as a peptide selected from the group consisting of Q475 and D679, the peptide has a sequence encoded by the construct using the universal code as is known in the art.

Protein Production & Purification:

This section details a protocol that we have used to produce the kinesin protein fragments detailed in Table II. Variations for particular kinesins are noted in the protocol. For many of the examples (Chromokinesin, Kin2), the protocols are quite similar. However, we have found that modifications to the protocol are preferred in certain cases. For example, for MCAK, the PIPES-based buffers were not suitable for production of active proteins, and therefore the success of Tris-based buffers were discovered.

Expression Protocol:

Typical culture volume for a preparation is 1-2 liters, with each 500 ml of culture being contained in a 2 liter flask to promote aeration. Typical culture media is LB medium with 10 ppm antifoam. Alternatively, TB medium is also suitable. Media is inoculated in the morning with a single fresh colony of bacterial cells (for example, *E. coli* strain BL21(DE3)plysS) harboring an expression plasmid (for example, those plasmids described above). For all kinesins, cultures are grown at 37°C with shaking until OD_{600} reaches about 0.8 at which point cultures continue to shake at room temperature for about 30-45 minutes. To provide a pre-induction sample, 500 µl of culture is spun down and frozen at -20°C at this point. To induce protein production, IPTG is added to 0.2 mM (or 0.5 mM for CENP-E and MCAK), and shaking is continued overnight. On the following morning (after 12-16 hours), another 500 µl sample is collected, spun down, and frozen at -20°C. The remainder of cells are harvested by centrifugation at 4°C for 30 minutes (for example, using a Beckman Allegra 6R Centrifuge at 3000 rpm or using a JLA 10 rotor in a Beckman Avanti J-25 centrifuge at 5000 rpm).

Purification Protocol:

The preferred buffers for each kinesin are described at the end of this section. From this point, all solutions are kept on ice and/or in a 4°C environment. Cell pellets are resuspended in lysis buffer supplemented with protease inhibitors (for example 1x concentrations of Complete EDTA-free protease inhibitors (Boehringer 1836 170)). 20 ml of lysis buffer is used for every 1 liter of culture. Dounce homogenization is conducted to ensure complete resuspension. At this point it is possible to freeze the cell suspension in liquid nitrogen and store it at -80°C. If cell suspension is frozen at this point, fresh DTT (and ATP for MCAK) are added upon thawing. Cells are lysed with a microfluidizer by running 2 passes, 7-8 cycles each at 80 psi. If cell suspension had been frozen, only 1 pass of 3 cycles is used. About 10 mls of extra lysis buffer is passed through the microfluidizer chamber to rinse it. Lysate is clarified by centrifugation (for example, for 45 minutes at 22,000 rpm in a JA25.50 rotor in a Beckman Avanti J-25 Centrifuge, or for 30-45 minutes at 30,000 rpm in a 45 Ti Rotor in a Beckman Optima LE-80K Ultracetnrifuge).

For MCAK, 0.5 ml of Ni-NTA resin (Qiagen 31014) is used for every 1 liter of culture. For all others, 1.5 ml of Ni-NTA resin is used for every 1 liter of culture. Resin is equilibrated with lysis buffer by washing 2 times with 15 ml of buffer without DTT and protease inhibitors. During these washes, resin is collected by spinning at 600-700 rpm for about 2 minutes in a bench-top

centrifuge. 100 µl of lysate is reserved before addition to the resin. Remainder of clarified lysate is added to the resin and incubated at 4°C for hour (20 minutes for MCAK) with rocking.

For Chromokinesin, Kin2, MKLP1, KSP and CENP-E, resin is collected by spinning at 600-700 rpm for about 2 minutes in a bench-top centrifuge. Supernatant is removed and a 100 µl sample is saved. Resin is resuspended in 5-10 ml lysis buffer with 0.1x protease inhibitors, and slurry is poured into a column For MCAK, lysate/resin slurry is directly poured into a column (for example, BioRad 1 cm ID EconoColumn), and flowthrough is collected and a 100 µl sample of flowthrough is reserved.

Column is then washed (using either gravity flow or a peristaltic pump at 1 ml/min) with 50 ml of lysate buffer. Column is then washed with 10 ml of wash buffer. Protein is eluted from column with 8 ml of elution buffer containing 0.1x protease inhibitors, and 1 ml fractions are collected. Fractions containing protein peak as measured by Bradford assay are pooled, and protein is diluted to 2 mg/ml with wash buffer with 0.1x protease inhibitors (for KSP, do not include Imidazole in wash buffer used for dilution). Aliquots are quick-frozen in liquid nitrogen and stored at -80°C.

Buffers Used in Purification Procedure

Chromokinesin, Kin2, MKLP-1, HSET, ATSV Buffers:

Lysis Buffer: 50 mM tris/HCl; 250 mM NaCl; 10 mM Imidazole; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 7.4.

Wash Buffer: 50 mM PIPES; 10% Sucrose; 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

Elution Buffer: 50 mM PIPES; 10% Sucrose; 300 mM Imidazole; 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

KSP Buffers:

Lysis Buffer: 50 mM tris/HCl; 250 mM NaCl; 10 mM Imidazole; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 7.4.

Wash Buffer: 50 mM PIPES; 10% Sucrose; 40mM Imidazole, 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

Elution Buffer: 50 mM PIPES; 10% Sucrose; 200 or 250 mM Imidazole; 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

CENP-E Buffers:

Lysis Buffer: 50 mM tris/HCl; 250 mM NaCl; 10 mM Imidazole; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; 0.1mM ATP, pH 7.4.

Wash Buffer: 50 mM PIPES; 10% Sucrose; 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

Elution Buffer: 50 mM PIPES; 10% Sucrose; 300 mM Imidazole; 100 mM NaCl; 2 mM $MgCl_2$; 1 mM EGTA; 1 mM DTT; pH 6.8 with NaOH.

MCAK Buffers:

Lysis Buffer: 50 mM tris/HCl; 50 mM NaCl; 10 mM Imidazole; 5 mM MgCl2; 1 mM EGTA; 1 mM DTT; 1mM ATP pH 6.8.

Wash Buffer: 50 mM tris/HCl; 50 mM NaCl; 50 mM Imidazole; 5mM MgCl2 1 mM EGTA; 1 mM DTT;1mM ATP, 20%sucrose; pH 6.8.

Elution Buffer: 50 mM tris/HCl; 50 mM NaCl; 100 mM Imidazole; 5mM MgCl2 1 mM EGTA; 1 mM DTT;1mM ATP; 20%sucrose; pH 6.8.

Results of Purification:

Successful application of this protocol is measured by the yield, purity and activity of the desired protein. Table III describes results using the protocol detailed above. We have assessed "activity" by the ability of the protein to hydrolyze ATP in a manner stimulated by microtubules. The motor domain of the kinesins is responsible for this enzymatic process. All of the constructs contain the motor domain, and differ in the amount of flanking sequence. We find that the character of the fragment can affect yield and purity (see Table III). We find that the purification conditions used can affect yield, purity and activity. The protocol above describes the most successful conditions, and Table III describes the outcome resulting from the preferred

protocol. There were also conditions tested that were not successful. For example, for KSP, elution buffer containing varying amounts of imidazole were tested. 50 mM and 100 mM imidazole-containing elution buffers failed to elute most protein, so yields were low. However, 200 mM and 250 mM imidazole-containing elution buffers resulted in high yields of active protein. As another example, for MCAK, the PIPES-based buffers were not suitable for production of active proteins as discussed above, therefore successful results were discovered with Tris-based buffers.

Table III: Production and Purification Results:

Kinesin	Construct Name	Residues Included	Production	Activity
Chromokinesin	K335	E2-K335	Expresses well	Low
Chromokinesin	Q475	E2-Q475	Expresses well	High
Chromokinesin	D679	E2-D679	Expresses well	High
Chromokinesin	FL1	E2-H1229	Does not express well	n/a
Kin2	P166	P166-A532	Expresses well	Yes
Kin2	H195	H195-A532	Expresses well	Yes
Kin2	FL2	M1-L679	Does not express well	n/a
MKLP1	E433	A4-E433	Expresses well	Yes
MKLP1	R494	A4-R494	Expresses well	Yes
MKLP1	E658	A4-E658	Does not express well	n/a
KSP	L360	M1-L360	Expresses well	Yes
KSP	K491	M1-K491	Expresses well	Yes
KSP	S553	M1-S553	Not as well as L360 and K491	n/a
CENP-E	M329	A2-M329	Expresses well, but relatively impure	Yes
CENP-E	T340	A2-T340	Expresses well	Yes
CENP-E	S405	A2-S405	Expresses well	Yes
CENP-E	V465	A2-V465	Expresses well, but relatively impure	Yes
CENP-E	T488	A2-T488	Expresses well, but relatively impure	Yes
MCAK	M1	R189-P617	Expresses well, low solubility	Low
MCAK	M2	P228-P617	Expresses well, low solubility	Low
MCAK	M3	H257-P617	Expresses well, low solubility	Moderate
MCAK	M4	R189-P660	Expresses well, low solubility	Low
MCAK	M5	P228-P660	Expresses well, low solubility	Low
MCAK	M6	H257-P660	Expresses well, low solubility	Moderate
MCAK	FL3	M3-Q725	Expresses well, low solubility	Low
Kid	A2N370	A2-N370	Expresses well	Not tested
Kid	A2M511	A2-M511	Expresses well	Not tested
HSET	K519	L72-K519	Low expression	Low
HSET	E152.2	E152-K519	Expresses well	Yes
HSET	Q151.3	Q151-K519	Expresses well	Yes
ATSV	Q353	G3-Q353	Expresses well	High
ATSV	M472	G3-M472	Expresses well	Low